Journal of Physics: Conference Series

PAPER • OPEN ACCESS

Measuring Cloud Base Height and Cloud Coverage using Elastic Multiangle Lidars at Pierre Auger Observatory

To cite this article: R Mussa *et al* 2022 *J. Phys.: Conf. Ser.* **2398** 012018

Measuring Cloud Base Height and Cloud Coverage using Elastic Multiangle Lidars at Pierre Auger Observatory

R Mussa[1], J Pallotta[2] for the Pierre Auger Collaboration[3]

[1] INFN Sezione di Torino, via Pietro Giuria 1, I-10125 Torino Italy
[2] Centro de Investigaciones en Láseres y Aplicaciones, UNIDEF (CITEDEF-CONICET), Villa Martelli, Buenos Aires, Argentina
[3] Observatorio Pierre Auger, Av. San Martin Norte 304, 5613 Malargue, Argentina
Full author list: https://www.auger.org/archive/authors_2022_07.html
E-mail: spokespersons@auger.org

Abstract. Cloud features above the Pierre Auger Observatory (Mendoza Province, Argentina) significantly affect the reconstruction of Extensive Air Showers. In this work, we present seasonal variations of cloud-base height, cloud coverage, and correlation between different sites using the information of elastic multiangle lidar data. This system locates the presence of clouds by measuring the spikes in the backscattered photons detected in the direction of the sweep performed during each lidar scan, outside the field of view (FOV) of the fluorescence detectors. Horizontal homogeneity should be assumed to translate these results to the full array. This ansatz is verified by a set of dedicated horizontal lidar shots performed for a few seconds every hour inside the FOV of the fluorescence detectors. Here we present the results for the period 2007 to 2016, using all the continuous lidar scans available in the lidar database. The analysis algorithm used for cloud retrieval has been upgraded and is based on a different concept than the previous one. How cloud parameters vary across seasons is investigated, and conclusions about cloud homogeneity across the Pierre Auger array are given.

1. Introduction

The Pierre Auger Observatory is the largest hybrid cosmic-ray detector in the world. Located in Argentina (35.32S, 69.30W, 1416 m a.s.l.), Malargüe, Mendoza, offers great atmospheric conditions to perform the measurements. Their facilities combine a large array of water Cherenkov detectors covering an area of 3000 km^2 and 24 fluorescence detectors (FDs) overlooking the surface array zone from different points of view. The FDs, located in the sites of Los Leones (LL), Los Morados (LM), Loma Amarilla (LA), and Coihueco (CO), are large telescopes aimed to detect the UV radiation produced by the interaction of the cosmic ray with nitrogen molecules. Their product accuracy relies on the atmospheric optical properties, which can be subject to its variations during the data taking. To face this task, the Observatory has an extensive program to monitor the atmosphere, measuring its properties during the fluorescence data taking. These measurements are meant to obtain the aerosol's extinction and cloud properties (base heights, coverage, and thickness).

One of the major obstacles related to extensive air shower observation by fluorescence detectors are clouds. A number of instruments are deployed to determine their parameters and used to apply proper quality cuts on cosmic ray analyses.
Central Laser Facility (CLF) and eXtra Laser Facility (XLF) [1], can return cloud heights as part of the aerosol profile determination for the application of the proper correction to air shower light profiles. Both facilities are based on a laser, firing scheduled pulses at the center of the array during the FD data-taking periods, measuring the scattered photons using each FDs telescope. The retrieval of cloud and aerosol parameters is done with the Data Normalized Analysis [1], which compares hourly laser light profiles to a reference profile registered on a very clear night. Also, infra-red (IR) cloud cameras at the FD sites provide measures of cloud coverage in the FD telescope fields of view. Studies measuring the cloud temperature using the IR data from the cameras were also done to derive cloud heights and coverage.
The multiangle-elastic backscatter lidars [2] co-located at each FD, are systems capable of producing aerosol profiles and cloud parameters in its line of sight. To accomplish this task, different scans are programmed during the FD data-taking periods based on the required output. In this sense, two continuous scans, on orthogonal planes, are performed every hour, to obtain cloud features, while discrete scans are used for aerosol retrieval. A typical image from two continuous scans is shown in Figure 1.

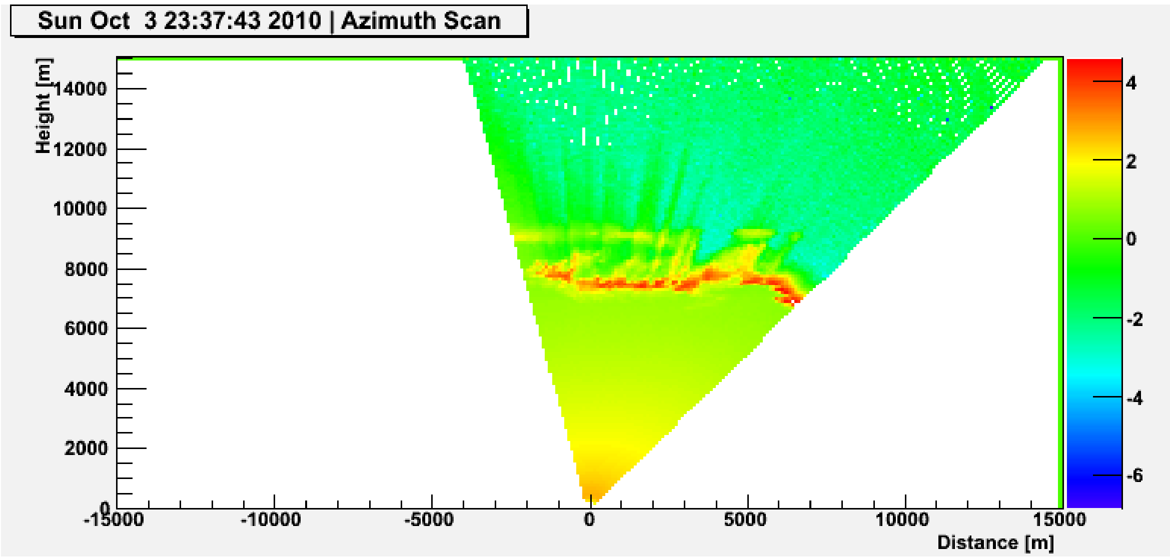

Figure 1. Example of multiangle continuous scan from Coihueco site. The image is obtained summing about 100 profiles (1000 laser shots each), acquired in ~ 10 minutes.

The continuous scans are analyzed to provide one measurement per hour of the lowest cloud base height and the highest fraction of the sky covered with clouds. To avoid interference with the FD, the lidars fire laser pulses outside the field of view of the FD. Hourly horizontal shots towards the pampa are used to measure the aerosol attenuation at ground level and the horizontal homogeneity, vetoing the acquisition of cosmic ray events for a few seconds every hour.
In this work, we report on the development of a new and more robust algorithm for the characterization of cloud parameters, based on a dynamic threshold instead of the differentials method used in the previous analysis. Site-to-site comparisons are made in order to study cloud homogeneity and to quantify with existing data the probability of extrapolating measurements done in the edges of the array to the whole area of the Observatory.

2. Lidar cloud detection algorithm

Different cloud base/top height definitions can be found in the literature, and there is not a single algorithm suitable for all clouds [3], being also strongly influenced by the measuring instrument [4]. The US National Weather Service defines cloud base height when the backscattered signal obtained using a rotating-beam ceilometer (RBC) reaches its maximum value [5]. This is not a convenient definition because the cloud base height is located inside the cloud. [4] defines cloud-base altitude as the altitude above which hydrometeors exist and can be detected to be liquid droplets, ice particles, or rain.

In the new algorithm developed, no cloud/aerosol plumes are discriminated, and only the increased backscatter layer in the elastic-lidar signals is detected (also called “feature” [6]). Once a feature is detected, different methods like [7] can be applied for cloud classification, and aerosol discrimination by inspecting the ratio of different parts of the lidar signal intensity across the feature. Most used algorithms are the core differential-based methods that rely on the differential of the range-corrected lidar signals (RCLS), dealing with the noise (especially at higher ranges) to avoid false positives. These methods deal with the fact that the intensity of the elastic lidar signals decreases monotonically with height unless a cloud/plume is encountered, where the signal level increases significantly. Based on this behavior, the methods aim to detect the change in the slope of the RCLS, which is carried out by the first derivative (dP/dz) of the elastic lidar signal. These slope changes are detected by setting a threshold to the derivative or by analyzing the zero-crossing to obtain their boundaries. Derivative-based methods like [4] reported measurements of the cloud base height and vertical extent by studying the vertical derivative of the raw lidar signal.

All these algorithms need input parameters that have to be set properly to avoid any false positives and reject small backscatter enhancements caused by noise. Also, their configuration depends on the lidar features, and no universal algorithm/parameters are reported capable of working in any elastic lidar hardware and atmospheric scenario. The new algorithm is based on a dynamic threshold to detect the layer, showing excellent results, even at UV wavelength, like the one used by the elastic-multiangle lidars of the Pierre Auger Observatory.

3. New multiangle-lidar cloud database inter-site intercomparison

The existing data on cross-correlations between sites are useful to compensate for periods during which one or more instruments were not operating, and allow checking the horizontal homogeneity of the highland where the Observatory is located. Significant differences are expected due to the presence of the Andes mountain range west of the array, and related turbulence due to the dominant winds from the Pacific Ocean. Data can be compared on two observables: cloud base height and cloud coverage. The cloud base height can change as multiple layers at different altitudes are not always extending through the whole array. Elastic lidars allow us to observe multiple layers of clouds, but only the base height of the lowest one is currently saved in our database.

3.1. Cloud Base Height

The plot in Figure 2 shows the cross-correlation between the cloud base heights (above sea level) measured in the sites of Coihueco (Y-axis) and Loma Amarilla (X-axis). The first two bins on each axis are filled by events when one of the two sites does not observe any cloud. The fraction of events when both sites are cloudless ranges between 30 and 40% across the seasons.

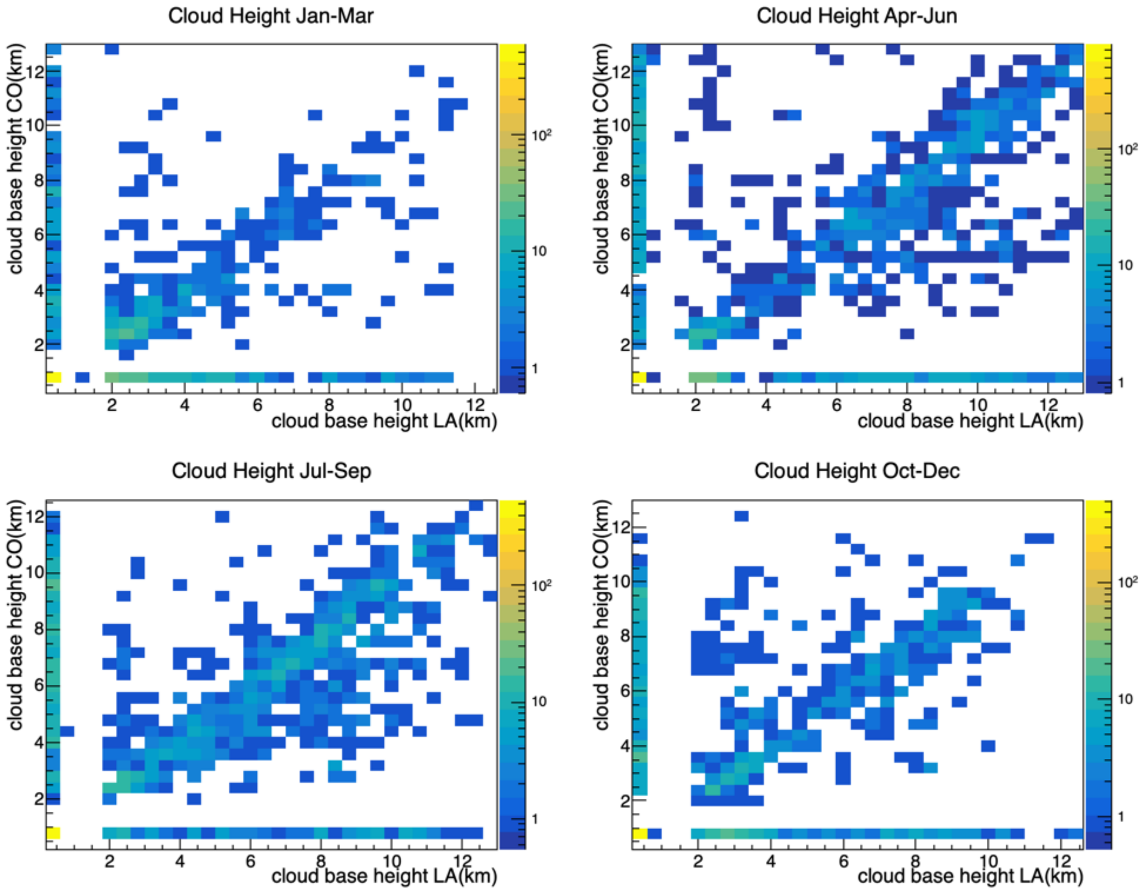

Figure 2. Cloud base height Cross-correlation between Coihueco (Y) and Loma Amarilla (X)

The scatter plots show an evident clustering of events on the diagonal, at all altitudes, consistent with a uniform distribution of cloud layers. Shallow cloud layers are more likely in Loma Amarilla and appear as a vertical band between 2 and 4 km. The distribution of medium-altitude clouds is asymmetric in autumn and winter, with a prevalence of higher clouds going inland.

In Figure 3 we compare the seasonal variations of the cloud height distributions in the two sites during the nights when both lidars observe clouds. During the austral summer, when storms are more frequent, large cumulus clouds with low base height are dominant. From spring to autumn, high layers (cirri) are observed up to 12 km altitude. A gap is present at about 4km asl, probably due to the Andes mountain range. The base height distributions are quite similar when clouds are observed in both sites, but differ significantly (in summer and autumn) when only one of the two sites is cloudy. Thunderstorm clouds are more frequent above the sites of Loma Amarilla and Los Morados.

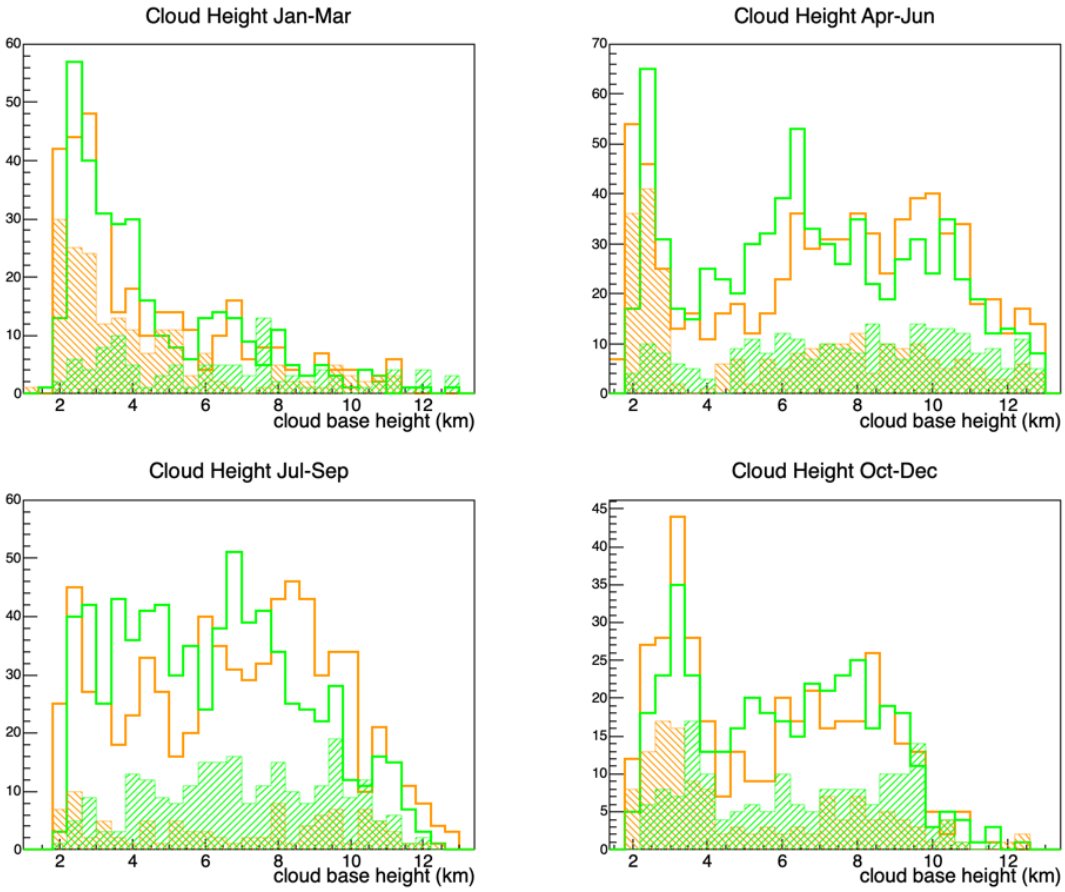

Figure 3. Cloud base height (a.s.l.) seasonal variation in Coihueco (green) and Loma Amarilla (orange). The line (hatched) histograms are when the other site sees (no) clouds.

In order to quantify the level of inhomogeneity, we plot the cumulative probability distribution function as a function of the cut on the difference between the measured cloud heights. The first bin also includes events with no clouds on both sites (30 to 40 percent of the total). The overall normalization of the fraction includes events when clouds are detected only in one of the two sites. Figure 4 shows the intercomparison of all the pairs of sites.

3.2. Cloud Coverage

Every hour, each lidar uses two continuous scans to estimate the cloud coverage above each FD site. It is computed by doing the rate between the number of profiles with detected clouds divided by the total number of profiles acquired in the scan. It is important to notice that if multiple layers are detected, only the minimum base height is stored, but the cloud coverage is measured by summing over all layers. In all the sites, the distribution of values of cloud coverage has two peaks at the two extreme values, 0 and 1. The cross-correlation between the two sites gives us another indication of possible inhomogeneity in the distribution of clouds across the array. As an example, Figure 5 compares the coverages in Los Leones and Coihueco. Typically, cloud coverage values cluster either close to 0 or close to 100 percent. Partial cloud coverage on one site is rarely correlated to similar cloud coverage on the other site.
As for the study of the cloud heights, we can parametrize the asymmetry in the distribution of cloud coverages, as the fraction of events above a certain value of the difference between its values in the two sites.

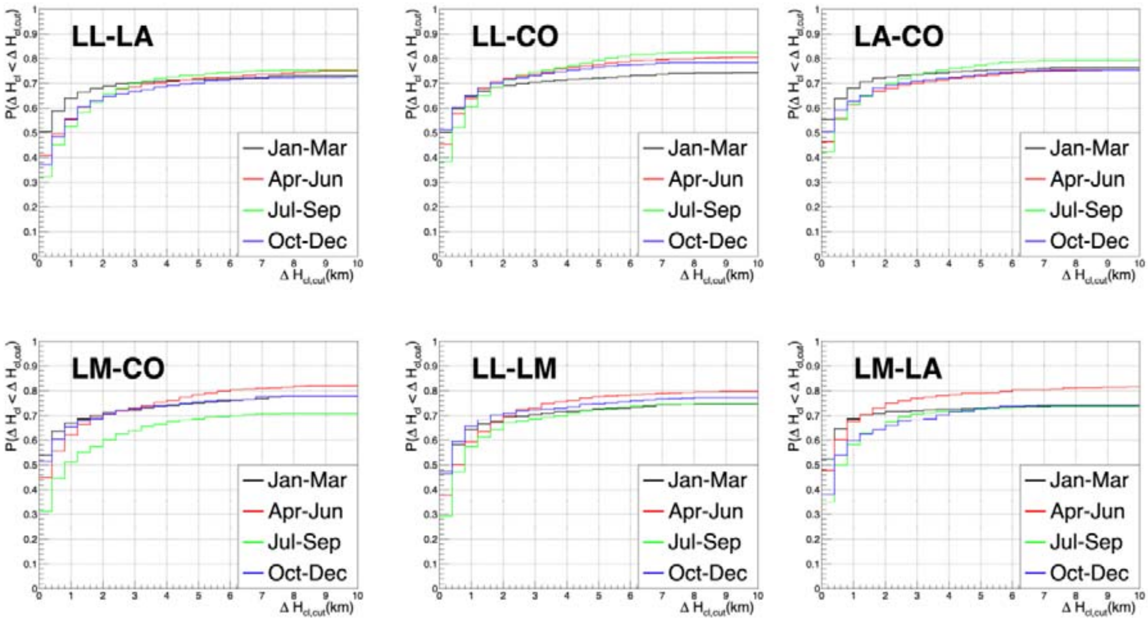

Figure 4. Cloud base height homogeneity as a function of the cut on the difference between the measured height.

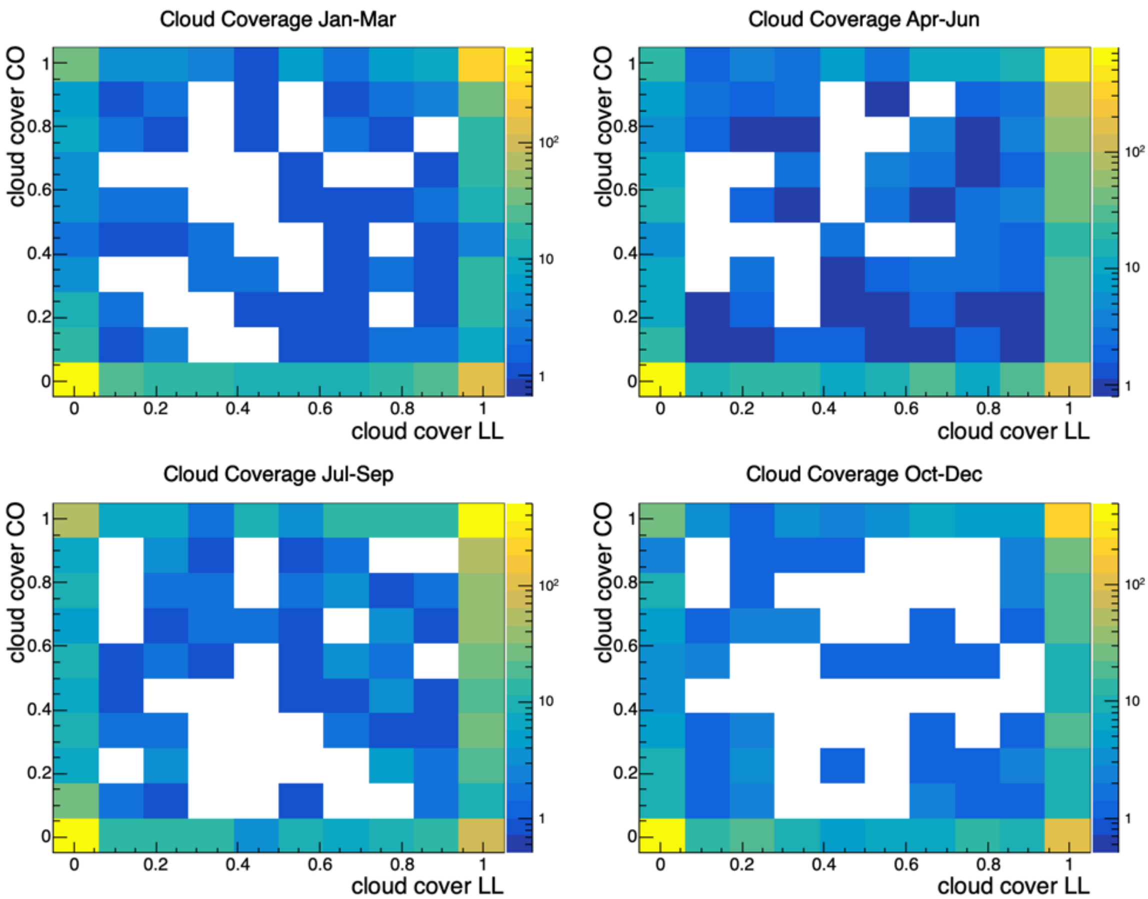

Figure 5. Cloud coverage cross-correlation between Coihueco and Los Leones.

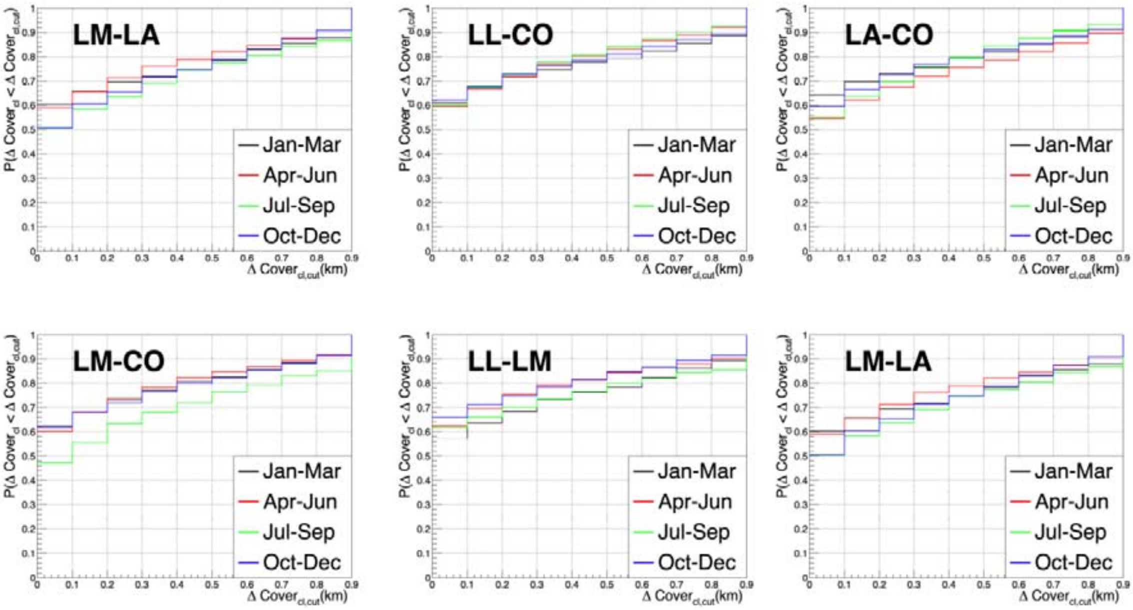

Figure 6. Cloud coverage homogeneity: cumulative distribution function of the cut on the difference between the measured values.

4. Conclusions

The elastic lidar network is a key component of the atmospheric monitoring system of the Pierre Auger Observatory, which provides real-time information on the cloud pattern above each FD site, during regular data-taking operations. The semi-automatic cloud retrieval algorithm, based on the method of the derivatives, has recently been upgraded to provide a more robust, fully automated extraction of cloud parameters. As two of the four elastic lidars were forced to stop operations in 2013 (Los Morados) and 2015 (Los Leones), we have to rely more heavily on the horizontal homogeneity of the cloud layers to extrapolate the information harvested above each site to the inner part of the array where the showers are observed. This report focuses on the asymmetries in cloud height and cloud coverage. Further studies are ongoing to correlate these observables to measurements provided by other instruments such as CLF/XLF, cloud cameras, weather stations, and satellites.

References

[1] The Pierre Auger Collaboration, 2013, *JINST* 8, 04009
[2] BenZvi S. et al., 2007, NIM A574, 171-184
[3] Platt C et al., 1994, *Bull. Amer. Meteor. Soc*. vol 75, 1635–1645
[4] Pal S et. al., 1992, *Appl. Opt.* vol 31, 1488–1494
[5] Eberhard W 1986 *J. Atmos. Oceanic Technol.* vol 3, 499–512
[6] Vaughan M. et. al., 2009, *Journal of Atmospheric and Oceanic Technology* vol 26, 2034
[7] Wang Z et. al., 2001, *Journal of Applied Meteorology* vol 40, 1665-1682